



Pratiques

Linguistique, littérature, didactique

169-170 | 2016

Enseignement/apprentissage de la langue, des textes
et des discours. 40 ans de *Pratiques*

L'apprentissage de la lecture (du comportement aux corrélats neuronaux) : un bilan de 30 ans de recherche

Learning to Read (from Brain to Behavior): A Survey of 30 Years of Research

Liliane Sprenger-Charolles



Édition électronique

URL : <http://journals.openedition.org/pratiques/2969>

DOI : 10.4000/pratiques.2969

ISSN : 2425-2042

Éditeur

Centre de recherche sur les médiations (CREM)

Référence électronique

Liliane Sprenger-Charolles, « L'apprentissage de la lecture (du comportement aux corrélats neuronaux) : un bilan de 30 ans de recherche », *Pratiques* [En ligne], 169-170 | 2016, mis en ligne le 30 juin 2016, consulté le 01 mai 2019. URL : <http://journals.openedition.org/pratiques/2969> ; DOI : 10.4000/pratiques.2969

Ce document a été généré automatiquement le 1 mai 2019.

© Tous droits réservés

L'apprentissage de la lecture (du comportement aux corrélats neuronaux) : un bilan de 30 ans de recherche

Learning to Read (from Brain to Behavior): A Survey of 30 Years of Research

Liliane Sprenger-Charolles

- 1 Cet article a pour objectif de dresser un bilan de plus de 30 ans de recherches sur l'apprentissage de la lecture. La première partie présente des résultats sur trois points essentiels à la compréhension des mécanismes en jeu dans la lecture et son apprentissage : d'une part, les capacités expliquant la compréhension écrite ; d'autre part, le rôle joué par le contexte dans l'identification des mots écrits et, enfin, les indicateurs de la mise en place d'automatismes dans cette identification. La seconde partie porte sur l'incidence de la transparence de l'orthographe et le rôle de la phonologie dans le développement des capacités d'identification des mots écrits. La conclusion comporte un résumé des acquis de la recherche accompagné par un cadre explicatif, ainsi que les principales implications pour les politiques éducatives.

Quelques éléments essentiels à la compréhension de certains débats actuels

1. Quelles capacités expliquent la compréhension écrite et quel rôle joue le contexte dans l'identification des mots écrits ?

1.1. Quelles capacités expliquent la compréhension écrite ?

- 2 Selon le *modèle simple de la lecture* proposé par W. A. Hoover et P. B. Gough (1990), la compréhension écrite est le produit de deux habiletés indépendantes : les habiletés de compréhension orale et celles dites de *décodage*. Depuis quelques années, ce modèle a été

précisé. En particulier, ce qui était dénommé *décodage* dans le modèle original a été spécifié et ce terme est utilisé maintenant uniquement pour désigner la procédure sublexicale (ou *procédure phonologique*), évaluée par la lecture de mots inventés (des pseudomots), qui ne peuvent pas être identifiés ou reconnus par une procédure lexicale.

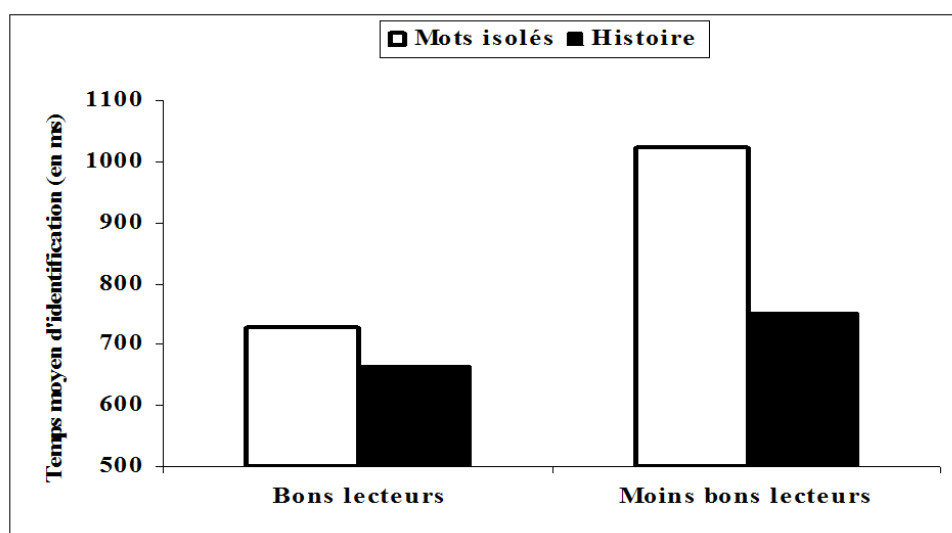
- 3 Sur la base de l'hypothèse que de bonnes capacités de décodage sont cruciales pour la compréhension écrite, une étude française récente a examiné la contribution du décodage et de la compréhension orale à cette capacité (Gentaz, Sprenger-Charolles & Theurel, 2015). Cette étude a pris en compte 392 enfants français de 1^{ère} primaire scolarisés dans des zones d'éducation prioritaire. Ce groupe a été divisé en trois sous-groupes selon le niveau de décodage des enfants, évalué par la lecture de pseudomots. Sur la base de l'écart-type¹ calculé par rapport à la moyenne du groupe, 267 enfants ont des scores de décodage moyens (entre +1 ET et -1 ET), 62 des scores faibles (inférieurs à 1 ET) et 63 de bons scores (supérieurs à 1 ET), soit environ 70 % dans le groupe des moyens, et 15 % dans chacun des deux groupes extrêmes.
- 4 Trois résultats importants ressortent de cette étude. Tout d'abord, le niveau de compréhension écrite est expliqué principalement par le niveau de décodage dans le groupe des décodeurs faibles et moyens alors que, chez les bons décodeurs, il est surtout expliqué par le niveau de compréhension orale. Ensuite, le pourcentage d'enfants ayant un problème de compréhension écrite est plus élevé dans le groupe de faibles décodeurs (55 %) que dans les deux autres groupes (7 % chez les décodeurs moyens et 0 % chez les bons décodeurs). Enfin, seulement 6 enfants (1,5 %) ont un déficit de compréhension écrite avec des capacités intactes en décodage, en compréhension orale et en vocabulaire.
- 5 Ces résultats reproduisent ceux d'une étude anglaise récente (cf. Spencer, Quinn & Wagner, 2014) et remettent en question la spécificité des *pauvres compreneurs* qui auraient un bon niveau de décodage (Oakhill, Cain & Bryant, 2003). En effet, ils signalent qu'une mauvaise compréhension de l'écrit s'explique, comme cela avait déjà été montré, par des déficits soit de décodage (Byrne, Freeboby & Gates, 1992), soit de compréhension orale (Juel, 1988 ; Stothard & Hulme, 1992). Plus généralement, le niveau de compréhension écrite est surtout expliqué par le niveau de décodage au début de l'apprentissage de la lecture alors que, dans les étapes ultérieures, il est surtout expliqué par le niveau de compréhension de l'oral (Keenan, Betjemann & Olson, 2008). Ainsi, chez des adultes qui savent lire, les corrélations entre compréhension orale et écrite sont très élevées (Gernsbacher, Varner & Faust, 1990). En d'autres termes, ceux qui comprennent bien à l'oral comprennent également bien à l'écrit et vice versa.

1.2. Quel rôle joue le contexte dans l'identification des mots écrits ?

- 6 Une autre question est celle du rôle joué par le contexte dans l'identification de mots écrits. La figure 1 reproduit les résultats d'une étude princeps dans laquelle des enfants de 5^e primaire (CM2), bons et moins bons lecteurs, devaient lire des mots présentés en isolat ou dans le contexte d'une histoire (Perfetti, Goldman & Hogaboan, 1979). Deux résultats marquants sont à souligner. D'une part, le temps moyen d'identification des mots écrits (c'est-à-dire le délai qui s'écoule entre la présentation d'un mot sur un écran et le début de la réponse vocale du sujet) est beaucoup plus long chez les moins bons lecteurs que chez les bons lecteurs quand les mots sont présentés en isolat (environ 300 ms). D'autre part, la différence entre les deux conditions est forte chez les moins bons lecteurs (environ 270 ms) alors qu'elle est minime chez les bons lecteurs. Cette étude

princeps montre, comme je l'ai écrit il y a 30 ans dans un numéro de *Pratiques* (1986, p. 16), que « les lecteurs les moins habiles sont aussi les plus dépendants du contexte et que cette dépendance s'explique par la non automatisation de leurs stratégies² d'identification des mots écrits ». D'autres résultats à l'appui de cette conclusion ont été obtenus dans des nombreuses autres études (pour des synthèses, cf. Perfetti, 1985 ; Stanovich, 2000).

Figure 1. Temps d'identification de mots en isolat ou dans le contexte d'une histoire : enfants de 5^e primaire bons et moins bons lecteurs (d'après Perfetti et al., 1979)



1.3. En résumé

- 7 Les études présentées dans la première partie de cette section indiquent que les relations entre décodage, compréhension orale et compréhension écrite varient en fonction du niveau de décodage : le niveau de compréhension écrite est, chez le lecteur débutant, fortement lié au niveau de décodage, alors que, chez celui qui maîtrise le décodage, il est fortement lié au niveau de compréhension orale. Celles présentées dans la seconde partie ont mis en relief le fait que, chez de bons lecteurs, les procédures d'identification des mots écrits sont très rapides et largement indépendantes du contexte. Seule la maîtrise d'automatismes dans l'identification des mots écrits permet au lecteur débutant d'atteindre un niveau de compréhension de l'écrit égal à celui de sa compréhension de l'oral, en libérant ses ressources cognitives du poids d'un décodage lent et laborieux. Des données permettant de cerner comment les procédures d'identification des mots écrits s'automatisent progressivement sont présentées dans la section suivante.

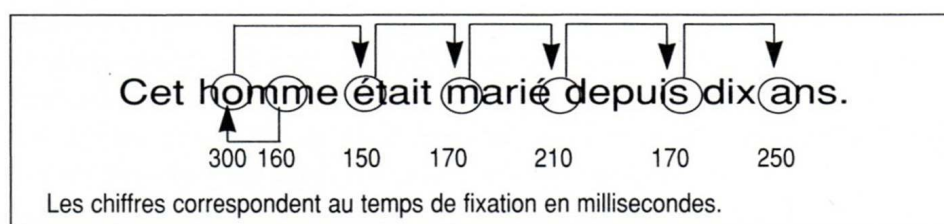
2. Comment ont été mis en évidence les automatismes dans les capacités d'identification des mots écrits ?

2.1. Au niveau de l'œil

- 8 Les yeux fournissent le signal qui déclenche l'activité de lecture et l'enregistrement des mouvements oculaires est un moyen d'observation de cette activité. L'information captée par les yeux est dirigée vers la rétine (notre plaque photographique), qui n'a pas un

pouvoir de discrimination identique dans toutes ses parties. En effet seule la fovéa assure la vision centrale, qui permet de distinguer avec précision les détails, en particulier les traits caractéristiques des lettres qui composent les mots. Or la fovéa n'occupe qu'une zone très limitée de la rétine et, au cours d'une fixation, il n'est possible de voir de façon nette qu'une petite partie de ce qui est lu. Cela permet de comprendre pourquoi le lecteur fixe pratiquement tous les mots d'un texte comme l'indique la figure 2 (d'après O'Regan & Lévy-Schoen, 1978). Dans cette figure, les parties encadrées indiquent les points de fixation et les flèches les saccades, c'est-à-dire les déplacements de l'œil d'un point à un autre. Les saccades sont en général des saccades de progression (celles en dessus de la ligne) et, plus rarement, des saccades de régression (celle en dessous de la ligne).

Figure 2. Mouvements oculaires pendant la lecture (d'après O'Regan & Lévy-Schoen, 1978)



- 9 Le comportement oculaire évolue avec l'expertise en lecture, comme l'indiquent les informations présentées dans le tableau 1 (d'après Starr & Rayner, 2001). En particulier la durée moyenne des fixations baisse, tout comme le nombre moyen de fixations pour 100 mots et la fréquence des régressions. Toutefois, même un adulte fixe pratiquement tous les mots quand il lit. De plus la durée des fixations est très courte, environ un quart de seconde chez les adultes, tout comme chez les enfants après trois à quatre ans d'apprentissage de la lecture. Ces résultats peuvent être pris à l'appui de l'hypothèse d'une automatisation progressive des procédures d'identification des mots écrits.

Tableau 1. Évolution des mouvements oculaires durant la lecture en fonction du grade³ (d'après Starr & Rayner, 2001)

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Adultes
Durée des fixations (ms)	355	306	286	266	255	240	233
Nombre de fixations pour 100 mots	191	151	131	121	117	106	94
Fréquence des régressions (%)	28	26	25	26	26	22	14

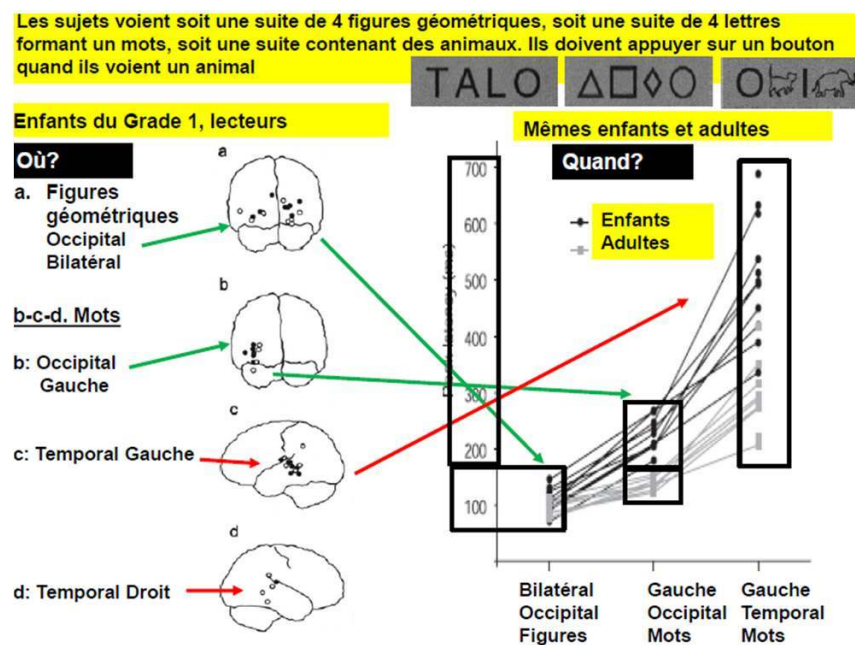
2.2. Données de neuro-imagerie

- 10 D'autres indicateurs d'une automatisation progressive des procédures d'identification des mots écrits ont été relevés plus récemment au niveau du cerveau. Par exemple, une étude

princeps a utilisé la magnétoencéphalographie pour voir où et quand se manifestent les activations consécutives à l'identification de mots écrits chez des enfants ayant appris à lire depuis un an (Parviainen et al., 2006). Ces enfants voyaient soit des items dits de *remplissage*, avec un animal, soit des items *tests* : un mot de 4 lettres ou une suite de 4 figures géométriques (triangle, carré, losange, cercle). Ils devaient appuyer sur un bouton quand ils voyaient un animal : leur attention n'était donc pas focalisée sur la lecture. Les analyses ont porté sur les réactions aux items *tests*.

- 11 Les résultats présentés dans la partie gauche de la figure 3 indiquent que les figures géométriques activent une aire occipitale bilatéralement (a) alors que cette aire est activée uniquement à gauche par les mots (b), ces derniers activant en plus l'aire temporale gauche chez la plupart des enfants (c). La partie droite de cette figure présente la temporalité de ces activations chez les enfants (gris foncé) ainsi que chez des adultes qui ont passé le même test (gris clair). Les activations pour les figures géométriques (a) se situent en moyenne autour de 100 ms chez les adultes, comme chez les enfants. Ce n'est pas le cas pour les mots (b et c). En effet, les activations correspondant à la réponse occipitale gauche (b), qui sont pratiquement dans la même fenêtre temporelle que celles observées pour les figures géométriques chez les adultes, sont un peu plus tardives chez les enfants (entre 200 et 300 ms). De plus, les activations correspondant à la réponse temporale gauche (c) sont encore plus tardives et varient entre sujets, adultes comme enfants, tout en étant toujours inférieures à la seconde.

Figure 3. Différences entre la perception de mots et de figures géométriques (d'après Parviainen et al., 2006)



- 12 Cette étude indique que, après quelques mois d'apprentissage de la lecture, les enfants, comme les adultes, identifient les mots écrits *en quelques millisecondes* ce qui est le signe d'un automatisme. Elle montre en plus que la lecture sollicite des zones spécifiques de l'hémisphère gauche du cerveau : l'une, située dans sa partie temporale, est connue pour traiter les aspects sonores du langage tandis que l'autre, située dans sa partie occipitale, est connue pour traiter les informations visuelles fines. Cette aire occipitale, qui a été recyclée pour traiter la forme visuelle des mots, est appelée *visual word form area* (VWFA ou

aire de la forme visuelle des mots). Elle va, au cours de l'apprentissage de la lecture, se connecter à d'autres aires de l'hémisphère gauche du cerveau dédiées au traitement du langage parlé : en particulier, des aires temporales et frontales. Avec l'amélioration du niveau de lecture, la VWFA devient de plus en plus activée par l'écrit et de moins en moins par d'autres stimulus et, si cette aire est détériorée, la lecture devient impossible (pour des synthèses, voir Dehaene, 2014 ; Dehaene & Cohen, 2011).

3. En résumé

- 13 Les résultats présentés dans la section 1 indiquent que, pour avoir un niveau de compréhension écrite ne dépendant que faiblement des capacités d'identification des mots écrits, il faut avoir automatisé ces dernières. Ceux présentés dans la section 2 montrent que cette automatisation s'explique non seulement par des modifications du système périphérique en charge de la lecture (nos yeux) mais également, au niveau central (le cerveau), par l'établissement, et la consolidation, de nouvelles connections neurales entre une aire visuelle spécifique (aire de la forme visuelle des mots) et les aires du langage. La partie suivante porte sur les facteurs – principalement linguistiques – qui facilitent la mise en place, et l'automatisation, des capacités d'identification des mots écrits.

Comment se mettent en place les procédures d'identification des mots écrits ?

- 14 Cette partie comporte quatre sections. La première développe un cadre de référence permettant de comprendre les problèmes rencontrés par un enfant quand il apprend à lire dans une écriture alphabétique. La seconde est centrée sur l'incidence de la transparence de l'orthographe sur les procédures d'identification des mots écrits. Cette section s'appuie sur des comparaisons interlangues, principalement entre enfants anglais, français, allemands et espagnols. La troisième, centrée sur des études effectuées avec des enfants français, présente des résultats concernant l'évolution des procédures d'identification des mots écrits. La quatrième porte sur le rôle joué par la phonologie dans l'apprentissage de la lecture. Une explication de l'ensemble des résultats présentés dans cet article est proposée en conclusion.

1. Un cadre de référence

- 15 Le lecteur débutant est confronté à plusieurs problèmes quand il apprend à lire dans une écriture alphabétique. Le premier vient de ce que, avant cet apprentissage, il n'a pas forcément accès de façon explicite aux unités de l'oral qui, dans ce type d'écriture, sont transcrites par des graphèmes, les phonèmes. En effet, les phonèmes, qui sont les unités distinctives minimales permettant de différencier – dans une langue donnée – deux mots (comme *bol* et *vol* en français) sont difficilement accessibles pour des raisons de coarticulation. C'est le cas pour les consonnes qui, comme l'indique leur nom, doivent toujours *sonner* avec autre chose, en l'occurrence une voyelle.
- 16 Un autre problème est lié au degré de consistance des correspondances entre graphèmes et phonèmes (CGP). Ainsi, alors que ces relations sont très consistantes en espagnol ou en

allemand, et un peu moins en français, l'anglais est clairement un *outlier*, comme l'indiquent les descriptions linguistiques et les données statistiques (Sprenger-Charolles & Colé, 2013, p. 58-68). Cela est particulièrement vrai pour les voyelles, un fait souligné par P. Delattre il y a plus de 50 ans (1965). Toutefois, en anglais, la prise en compte de la rime (la voyelle et les consonnes qui suivent) permet d'augmenter fortement cette consistance.

- 17 Dans une écriture alphabétique, le lecteur peut lire les mots écrits par deux procédures : soit une procédure lexicale, également appelée procédure orthographique ; soit une procédure sublexicale, procédure qui utilise les CGP et qui est également appelée procédure phonologique ou décodage (Coltheart, Rastle, Perry, Langdon & Ziegler, 2001 ; Ziegler, Perry & Zorzi, 2014). La première procédure est généralement évaluée par la lecture de mots fréquents irréguliers, qui ne peuvent pas être lus par la voie phonologique et la seconde par la lecture de pseudomots. La question est de savoir comment se mettent en place ces procédures de lecture au cours de l'apprentissage dans des écritures ayant des CGP plus ou moins consistantes.
- 18 Si le degré de consistance des CGP a un effet sur l'apprentissage de la lecture, des différences quantitatives devraient être relevées entre les apprentis-lecteurs espagnols, allemands, français et anglais. Les Espagnols et les Allemands devraient en effet apprendre à lire mieux et plus vite que les Français, qui eux-mêmes devraient apprendre à lire mieux et plus vite que les Anglais. De plus, si les procédures mises en œuvre par les lecteurs débutants sont fonction des caractéristiques spécifiques de chaque langue, des différences qualitatives devraient également être relevées. Ainsi, apprendre à lire dans une orthographe transparente (en espagnol ou en allemand) devrait conduire à une plus grande utilisation de la voie phonologique de lecture, l'inverse étant attendu quand l'orthographe est, comme celle de l'anglais, peu consistante.

2. Résultats d'études interlangues

- 19 Comme attendu, les études interlangues signalent que le niveau en lecture dépend du degré de transparence de l'orthographe. En effet, les scores les plus faibles s'observent chez les anglophones et les meilleurs chez les hispanophones. C'est ce qui ressort, par exemple, d'une vaste étude comparative qui a impliqué des enfants de différents pays d'Europe : Allemagne, Espagne, Finlande, France, Grande-Bretagne, Italie, Pays-Bas... (Seymour, Aro, Erskine et al., 2003). Excepté les Anglais et les Français, qui ont également été examinés après 2 ans d'apprentissage de la lecture, les autres enfants n'ont été vus qu'après une année d'apprentissage. Les anglophones de 1^{re} année sont les plus jeunes (5 ans ½) et ceux de 2^e année ont approximativement le même âge (6 ans ½) que les enfants français et espagnols de 1^{re} année. Ces enfants ont eu à lire des mots fréquents (noms, adjectifs ou verbes : par exemple, *voir* et *femme* en français ou *high* et *boy* en anglais) et des mots de fonction (par exemple, *donc* et *alors* en français ou *them* en anglais) et des pseudomots.
- 20 À niveau scolaire équivalent (1^{re} primaire), les scores des anglophones sont plus faibles que ceux des hispanophones et des francophones. À âge équivalent (6 ans ½), bien qu'ayant bénéficié d'une année supplémentaire d'apprentissage de la lecture, les scores des anglophones sont encore inférieurs à ceux des hispanophones, mais plus à ceux des francophones. Ces résultats sont d'autant plus remarquables que les anglophones de cette

étude ont un niveau de lecture supérieur aux normes nationales et sont issus de milieux socioéconomiques plutôt favorisés.

- 21 Les facteurs linguistiques induisent également des différences qualitatives. Ainsi, les anglophones utilisent davantage la procédure lexicale, probablement pour compenser la difficulté de mise en œuvre du décodage dans leur langue. C'est ce que suggère le fait que les différences les plus fortes à leur détriment sont relevées en lecture de pseudomots, et non en lecture de mots. Ce résultat est généralement relevé dans les études interlangues en comparaison avec des enfants français, par exemple (Bruck, Genesee & Caravolas, 1997 ; Goswami, Gombert & Barrera, 1997). Dans celles impliquant des germanophones, la présence de corrélations très élevées entre lecture de mots et de pseudomots ainsi que l'absence d'effets facilitateurs de la fréquence signalent que ces derniers (mais pas les anglophones) ont essentiellement recours à la procédure phonologique (Frith, Wimmer & Landerl, 1998 ; Wimmer & Goswami, 1994 ; voir pour des résultats similaires avec des enfants francophones vus en milieu de 1^{re} primaire, Sprenger-Charolles, Siegel & Bonnet, 1998b). La moindre consistance des CGP en anglais favorise également un plus grand recours à des unités plus larges que le graphème, en particulier la rime des mots (cf. Goswami et al., 1998 ; Goswami, Ziegler, Dalton & Schneider, 2003) parce que, dans cette langue, la prononciation des voyelles dépend largement des consonnes qui suivent. Ce n'est pas le cas dans des langues qui ont des voyelles clairement articulées, comme en espagnol (Goswami et al., 1998), ou en allemand (Goswami et al., 2003), entre autres.

3. Développement des procédures de lecture en français

3.1. Données statistiques sur l'orthographe du français

- 22 Dans la suite du travail de N. Catach (1980), la consistance des correspondances graphème-phonème (CGP, utilisées pour lire) et phonème-graphème (CPG, utilisées pour écrire) a été examinée en français en tenant compte de la morphologie (principalement des flexions) dans le lexique écrit adressé à l'enfant (Peereman, Sprenger-Charolles & Messaoud-Galusi, 2013). Le corpus comprend les 10 000 mots les plus fréquents de la base de données *Manulex* construite à partir de 54 manuels du primaire (Lété, Sprenger-Charolles & Colé, 2004). Cette base inclut environ 2 000 000 d'occurrences, 98 % étant comprises dans les 10 000 formes les plus fréquentes sélectionnées pour *Manulex-Morpho*. Les calculs effectués ont pour objectif de mettre en relief, par exemple, la fréquence du graphème *s* par rapport aux autres graphèmes du français, tout comme la consistance des CGP pour ce graphème, à savoir, combien de fois il se lit /s/ comme dans *sol* par rapport au nombre de fois où il se lit /z/ comme dans *rose* et au nombre de fois où il est muet, comme dans *gris* ou *amis*. On a également calculé combien de fois ce graphème correspond à une marque morphologique de flexion (par exemple, le *s* du pluriel) ou de dérivation (par exemple, le *s* de *bois*, qui donne *boiserie*), par rapport au nombre de fois où ce n'est pas le cas (comme dans *alors*).
- 23 Le tableau 2, qui présente un résumé des principaux résultats, fait ressortir une forte asymétrie entre les CGP et les CPG, les premières étant plus consistantes que les secondes dans tous les cas (place dans le mot, prise en compte ou non de la morphologie). Cela s'explique par le fait que, par exemple, il n'y a qu'une façon de lire les graphèmes *o*, *au* et *eau* alors que pour écrire des mots contenant le phonème /o/, il faut choisir entre les différents allographes de ce phonème. Autre fait marquant : l'asymétrie de consistance des CGP comparativement aux CPG, qui est particulièrement marquée en fin de mot

quand la morphologie n'est pas prise en compte (12 %, soit 78 % pour les CGP et 66 % pour les CPG), diminue fortement quand elle l'est (5 %, soit 80 % versus 75 %). Et ce sont surtout les CPG qui bénéficient de cette prise en compte : le gain est de pratiquement 10 % (de 66 à 75 %), contre 2 % pour les CGP (de 78 à 80 %).

Tableau 2. Consistance des correspondances graphème-phonème (CGP) et phonème-graphème (CPG) en fonction de la position dans le mot et de la prise en compte des morpho-phonogrammes (d'après Peereman et al., 2013)

Position dans le mot	CGP (lecture)		CPG (Ecriture)	
	Sans morphologie	Avec morphologie	Sans morphologie	Avec morphologie
Début	96 %	96 %	87 %	87 %
Milieu	74 %	75 %	72 %	72 %
Fin	78 %	80 %	66 %	75 %

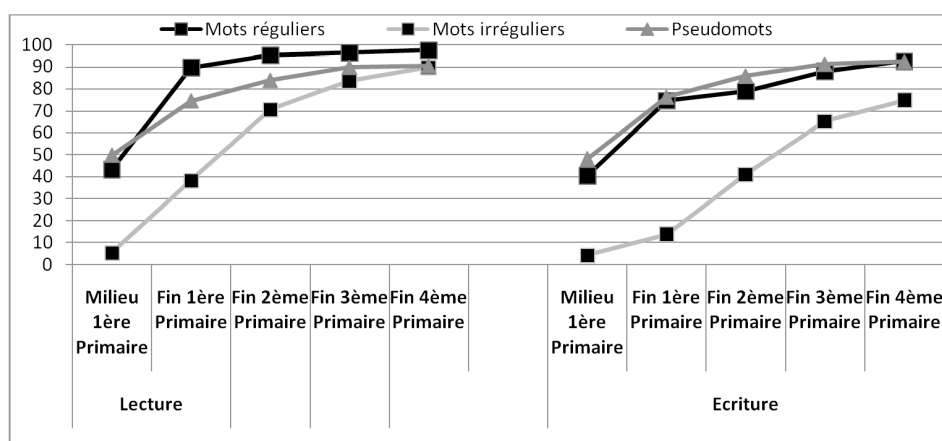
- 24 Étant donnée la régularité des CGP en français, les enfants français devraient, en principe, avoir principalement recours à la procédure phonologique en début d'apprentissage de la lecture, l'efficacité initiale de cette procédure devant prédire leur devenir en lecture. Cette hypothèse a été évaluée dans plusieurs études longitudinales dans lesquelles les enfants ont été suivis du milieu à la fin de la 1^{re} année du primaire (Sprenger-Charolles et al., 1998b) ou à la fin de la 4^{ème} année (Sprenger-Charolles, Siegel, Béchennec & Serniclaes, 2003), ainsi que du début de la grande section à la fin de la 2^e primaire (Sprenger-Charolles, Siegel & Béchennec, 1998a) à l'aide d'épreuves de lecture à haute voix et d'écriture sous dictée de mots fréquents, réguliers et irréguliers, et de pseudomots de même niveau de difficulté orthographique que les mots réguliers. Elle a également été évaluée dans certaines des études précédentes (Sprenger-Charolles et al., 1998a ; 2003) ainsi que dans d'autres (Grainger, Lété, Bertrand, Dufau & Ziegler, 2012) avec des épreuves de lecture silencieuse : épreuve de décision sémantique ou de décision lexicale contenant des pseudomots homophones de mots fréquents (rouge, naige...).

3.2. Résultats obtenus en lecture à haute voix et en écriture sous dictée

- 25 Si les enfants utilisent essentiellement la procédure phonologique au début de l'apprentissage, les mots réguliers doivent être lus aussi bien que les pseudomots et mieux que les mots irréguliers. C'est le résultat observé en milieu de 1^{re} primaire, les différences entre mots réguliers et pseudomots ne devenant significatives qu'en fin d'année (Sprenger-Charolles et al., 2003, cf. la figure 4). Toutefois, en fin de 1^{re} primaire, la procédure phonologique est encore utilisée et même plus fortement qu'auparavant. Ainsi, la différence entre les scores en lecture de mots réguliers et irréguliers augmente (de 38 à 51 %), tout comme la proportion d'erreurs de régularisation (*sept* lu comme *septembre*) qui est 5 fois plus forte en fin de 1^{re} primaire qu'au milieu de la même année (27 % du nombre total d'erreurs contre 5 %, Sprenger-Charolles et al., 1998b). C'est entre ces deux périodes que se produisent les changements les plus notables, tous les scores s'améliorant alors fortement, mais davantage ceux pour les mots réguliers que ceux pour

les mots irréguliers et les pseudomots. Inversement les mots réguliers ne sont jamais mieux écrits que les pseudomots. En dépit de ces différences entre lecture et écriture, il y a de fortes similarités entre ces deux modalités, comme le signalent les corrélations : 0,85 pour les pseudomots, 0,80 pour les mots réguliers et 0,72 pour les mots irréguliers en fin d'année (Sprenger-Charolles et al., 1998b ; voir aussi Eme & Golder, 2005).

Figure 4. Pourcentage de réponses correctes en lecture (à gauche) et en écriture (à droite) de mots et de pseudomots : du milieu de la 1^{re} année du primaire à la fin de la 4^{ème} année (CP au CM1). D'après Sprenger-Charolles et al. (2003)



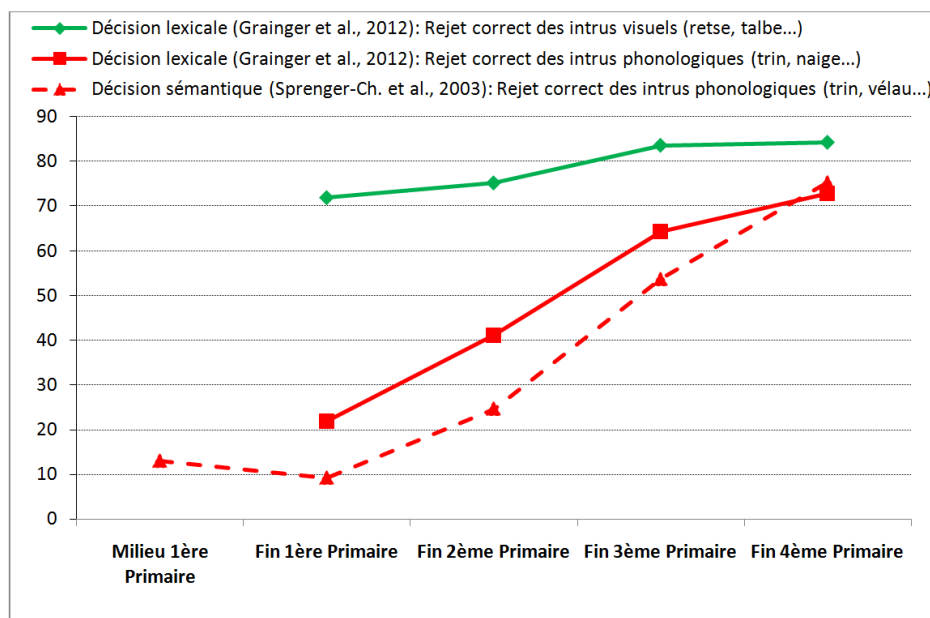
- 26 Ces résultats indiquent que, au début de l'apprentissage, les enfants utilisent principalement la procédure phonologique, le poids de cette procédure augmentant entre le milieu et la fin de la 1^{re} primaire, alors même que la procédure lexicale se met en place⁴. Dans les études précédentes, les enfants devaient lire à haute voix, ce qui a pu conduire à surestimer le rôle de la phonologie. Cela a été vérifié à l'aide d'épreuves de lecture silencieuse.

3.3. Résultats obtenus en lecture silencieuse

- 27 L'étude transversale⁵ de J. Grainger et al. (2012) a pris en compte des enfants des 5 niveaux du primaire qui ont eu à effectuer une tâche de décision lexicale avec des pseudomots-homophones et des pseudomots intrus visuels (construits en transposant des consonnes dans un mot : *retse*, *femre*). Les résultats pour les 4 premiers niveaux sont présentés dans la figure 5. L'effet de la proximité visuelle est faible (70 % de rejets corrects des intrus visuels dès la 1^{re} primaire) et il n'évolue que peu avec l'âge. Par contre, l'effet de l'homophonie, qui est très fort au début de l'apprentissage (20 % de rejets corrects des pseudomots-homophones en fin de 1^{re} primaire et 40 % un an plus tard), baisse avec l'âge. Comme l'indique également cette figure, et comme noté par J. Grainger et al. (2012), les résultats relevés pour les pseudomots-homophones correspondent de façon remarquable à ceux rapportés dans l'étude longitudinale de L. Sprenger-Charolles et al. (2003) qui a utilisé une tâche de décision sémantique (après la question *Est-ce que c'est une couleur*, l'enfant voyait soit un mot correct, comme *jaune*, soit un pseudomots-homophone, comme *rouje* ou un intrus visuel, comme *rouqe*). Cette étude comportait une épreuve complémentaire de choix orthographique dans laquelle l'enfant devait désigner le mot correct dans une suite de mots dans laquelle il était présenté en même temps que les intrus phonologiques et visuels (*vélo*, *vélau*, *véla* ; *loup*, *lou*, *louq...*). L'objectif de cette épreuve était de vérifier que les enfants connaissaient bien l'orthographe des mots

utilisés dans l'épreuve de décision sémantique. C'est le cas à partir de la fin de la 2^{de} année, le nombre moyen de réponses correctes étant alors de 86 %. Or à cette époque les enfants ne rejettent correctement que 25 % des intrus phonologiques dans l'épreuve de décision sémantique. Ce résultat signale que, y compris chez des enfants connaissant bien l'orthographe des mots utilisés, l'activation du code phonologique des mots écrits est difficile à inhiber, même quand cette activation n'est pas sollicitée par la tâche.

Figure 5. Évolution des réponses correctes entre le milieu de la 1^{re} année du primaire et la fin de la 4^e année (CP-CM1) dans une épreuve de décision lexicale (données transversales, adapté de Grainger et al., 2012) ou de décision sémantique (données longitudinales, adapté de Sprenger-Charolles et al., 2003)



4. Quel rôle joue la phonologie dans l'apprentissage de la lecture ?

- 28 Une hypothèse forte avancée par D. L. Share en 1995 (voir aussi Ziegler et al., 2014) est que la maîtrise du décodage est le sine qua non de l'apprentissage de la lecture. Les enfants qui ont le plus recours au décodage en début d'apprentissage, c'est-à-dire ceux qui lisent et écrivent bien les pseudomots ou encore ceux qui font souvent des erreurs phonologiques en lecture à haute voix (erreurs de régularisation: *sept* lu comme *septembre*) ou en lecture silencieuse (acceptation erronée de pseudomots homophones de mots) sont ceux qui progressent le plus (études anglophones : Byrne et al., 1992 ; Jorm, Share, MacLean & Matthews, 1984 ; études francophones : Sprenger-Charolles et al., 1998a). Par exemple, dans l'étude francophone citée, les enfants qui se révèlent être des experts en orthographe à 8 ans (ceux qui ne faisaient alors aucune erreur dans l'épreuve de choix orthographique) étaient ceux qui, un an avant, acceptaient le plus souvent des pseudomots-homophones comme étant des mots (dans 95 % des cas), ce qui signale qu'ils n'arrivaient pas à inhiber l'information phonologique en lecture silencieuse. Ces enfants étaient également ceux qui, à la même époque, utilisaient le mieux le décodage en lecture à haute voix ou en écriture sous dictée. À 5 ans, ces enfants avaient également les scores les plus élevés dans une tâche d'analyse phonémique.

Possibles explications et implications pour l'éducation

- 29 Les résultats présentés dans cet article s'expliquent dans le cadre de l'hypothèse développée par D. L. Share (1995) et reprise par J. C. Ziegler et al. (2014) selon laquelle l'utilisation du décodage constitue un puissant mécanisme d'autoapprentissage. C'est actuellement la seule hypothèse qui permet d'intégrer la majeure partie des résultats de la recherche sur l'apprentissage de la lecture⁶. Si les enfants utilisent d'abord le décodage, des connexions vont progressivement se créer entre graphèmes et phonèmes, l'établissement de ces connexions dépendant de la consistance des CGP. C'est le résultat qui ressort des études interlangues, les scores de lecture les plus bas s'observant chez les anglophones, qui sont confrontés à une orthographe très opaque, et les plus hauts chez les hispanophones, qui sont confrontés à une orthographe très transparente (Seymour et al., 2003).
- 30 L'importance de la phonologie ressort également des résultats observés dans les études françaises qui ont utilisé des tâches de lecture silencieuse (Grainger et al., 2012 ; Sprenger-Charolles et al., 2003). Ces études montrent en effet que les enfants prennent un pseudomot-homophone pour un mot très souvent dans les débuts de l'apprentissage de la lecture. Toutefois la fréquence joue rapidement un rôle. Ainsi, dans les études longitudinales françaises, alors que, en milieu de 1^{re} primaire, les scores en lecture de mots réguliers et de pseudomots ne diffèrent pas et sont supérieurs à ceux relevés pour les mots irréguliers alors très faibles, en quelques mois, le tableau se modifie fortement. Les mots réguliers sont en effet, en fin de 1^{re} primaire, mieux lus que les pseudomots et les mots irréguliers. Cela provient de ce que les mots réguliers sont réguliers et fréquents. Ni la fréquence seule, ni la régularité seule suffisent, comme l'indique la progression plus faible relevée pour les mots irréguliers d'une part, et pour les pseudomots d'autre part, comparativement aux mots réguliers (Sprenger-Charolles et al., 1998b, 2003).
- 31 Les facteurs linguistiques induisent également des différences qualitatives. Ainsi, la procédure lexicale est davantage utilisée quand l'orthographe est opaque, probablement pour compenser les difficultés de mise en œuvre du décodage (Goswami et al. 1998, 2003 ; Seymour et al., 2003). Les facteurs linguistiques induisent aussi des différences dans les unités de lecture utilisées (voir pour une synthèse Ziegler & Goswami, 2005). Par exemple, l'anglais favorise une plus grande utilisation des rimes des mots parce que, dans cette langue la prononciation des voyelles dépend largement des consonnes qui suivent (Goswami et al. 1998, 2003). Toutefois, quelle que soit la langue, c'est la capacité de décodage qui détermine le succès de l'apprentissage de la lecture, les meilleurs décodeurs précoces étant ceux qui progressent le plus, y compris en lecture de mots irréguliers et en compréhension écrite (Byrne et al., 1992 ; Sprenger-Charolles et al., 2003).
- 32 Le cadre explicatif proposé permet également de comprendre pourquoi les prédicteurs les plus fiables de l'apprentissage de la lecture sont les capacités d'analyse phonémique, y compris dans les études qui ont contrôlé le niveau de pré-lecture, les preuves à l'appui d'une contribution précoce des capacités d'analyse d'unités plus larges (syllabe ou attaque-rime) étant faibles (voir pour une synthèse, Melby-Lervåg, Lyster, & Hulme, 2012 et pour une étude longitudinale récente avec des enfants français, Piquard-Kipffer & Sprenger-Charolles, 2013). Cela provient de ce que les capacités d'analyse phonémique sont cruciales au début de l'apprentissage de la lecture dans une écriture alphabétique, parce qu'elles permettent à l'enfant d'accéder au principe de ce type d'écriture.

- 33 Ce cadre explicatif permet également de comprendre pourquoi un enseignement utilisant la phonologie, et ce de façon systématique, est plus efficace que toutes les autres méthodes. Ainsi, les enfants exposés à ce type d'enseignement obtiennent des résultats supérieurs en décodage et en compréhension écrite à ceux des enfants qui ont bénéficié d'autres méthodes. De plus, cet enseignement est particulièrement bénéfique pour les enfants à risque de difficultés d'apprentissage de la lecture, que ce soit pour des raisons linguistiques ou sociologiques (pour une synthèse, voir Sprenger-Charolles & Colé, 2013, p. 238-248)⁷.
- 34 La finalité de la lecture est la compréhension de ce qui est lu. Cependant, comme le signalent les données présentées dans cet article, la plupart des difficultés d'apprentissage de la lecture s'expliquent non par de difficultés de compréhension, mais par des difficultés d'identification des mots écrits. C'est là un résultat essentiel pour une pédagogie fondée sur la preuve (Dehaene et al., 2011). En effet, dans la mesure où l'objectif principal de l'enseignement est de permettre à l'enfant de parvenir à comprendre ce qu'il lit de la même façon qu'il comprend ce qu'il entend, l'enseignant doit l'aider à développer des capacités d'identification des mots écrits rapides et précises. Il doit également, bien entendu, l'aider à améliorer sa compréhension orale. Cela requiert une formation correcte en linguistique, en particulier une bonne connaissance des correspondances graphème-phonème (basée sur des données de corpus et non sur quelques exemples ad hoc !) ainsi que d'un autre domaine dans lequel l'écrit diffère fortement de l'oral : la morphologie. Cette formation doit inclure une approche des composantes syntaxiques et sémantiques du langage impliquées dans la compréhension, composantes pour lesquelles les différences entre écrit et oral sont relativement peu importantes (question qui n'a pas été abordée dans le présent article). Cela requiert également une bonne connaissance des travaux de psycholinguistique, en particulier de ceux qui ont porté sur l'identification et la validation des moyens pouvant permettre aux enfants de bien apprendre à lire.

BIBLIOGRAPHIE

- BRAIBANT, J.-M. & GÉRARD, F.-M. (1996). « Savoir lire: question(s) de méthodes ? » *Bulletin de psychologie scolaire et d'orientation* 1, p. 7-45.
- BRUCK, M., GENESEE, F. & CARAVOLAS, M. (1997). « A cross linguistic study of early literacy acquisition ». In: Blachman, B. (ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention*. Mahwah : Lawrence Erlbaum, p. 145-162.
- BYRNE, B., FREEBODY, P. & GATES, A. (1992). « Longitudinal data on the relations of word-reading strategies to comprehension, reading time and phonemic awareness ». *Reading Research Quarterly* 27, p. 141-151.
- CATACH, N. (1980). *L'Orthographe française. Traité théorique et pratique*. Paris : Nathan.

- COLTHEART, M., RASTLE, K., PERRY, C., LANGDON, R. & ZIEGLER, J. C. (2001). « DRC: A dual route cascaded model of visual word recognition and reading aloud ». *Psychological Review* 108, p. 204-256.
- DEHAENE, S. (dir.), DEHAENE-LAMBERTZ, G., GENTAZ, E., HURON, C. & SPRENGER-CHAROLLES, L. (2011). *Apprendre à lire : des sciences cognitives à la salle de classe*, Paris : Odile Jacob.
- DEHAENE, S. (2014). « Reading in the brain revised and extended: Response to comments ». *Mind and Language* 29 (3), p. 320-335.
- DEHAENE, S. & COHEN, L. (2011). « The unique role of the visual word form area in reading ». *Trends in Cognitive Sciences* vol. 15, p. 254-262.
- DELATTRE, P. (1965). *Comparing the phonetic features of English, French, German and Spanish*. Heidelberg : Jumiuss Gross Verlag.
- EME, E. & GOLDER, C. (2005). « Word-reading and word-spelling styles of French beginners: Do all children learn to read and spell in the same way? ». *Reading and Writing: An Interdisciplinary Journal* 18, p. 157-188.
- FRITH, U., WIMMER, H., & LANDERL, K. (1998). « Differences in phonological recoding in German- and English-speaking children ». *Scientific Study of Reading* 2, p. 31-54.
- GENTAZ, E., SPRENGER-CHAROLLES, L. & THEUREL, A. (2015). « Differences in the predictors of reading comprehension in first graders from low socio-economic status families with either good or poor decoding skills ». *Plos One* 10(3), e0119581.
- GERNSBACHER, M. A., VARNER, K. R. & FAUST, M. E. (1990). « Investigating differences in general comprehension skills ». *Journal of Experimental Psychology: Learning, Memory and Cognition* 16, p. 430-445.
- GOIGOUX, R. (2000). « Apprendre à lire à l'école : les limites d'une approche idéovisuelle. *Psychologie Française* 45, p. 233-243.
- GOSWAMI, U., GOMBERT, J. E. & BARRERA, L. F. (1998). « Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French and Spanish ». *Applied Psycholinguistics* 19, p. 19-52.
- GOSWAMI, U., ZIEGLER, J. C., DALTON, L. & SCHNEIDER, W. (2003). « Nonword reading across orthographies: How flexible is the choice of reading units? ». *Applied Psycholinguistics* 24, p. 235-247.
- GRAINGER, J., LÉTÉ, B., BERTAND, D., DUFAU, S. & ZIEGLER, J. C. (2012). « Evidence for multiple routes in learning to read ». *Cognition* 123, p. 280-292.
- HOOVER, W. A. & GOUGH, P. B. (1990). « The simple view of reading ». *Reading and Writing: An Interdisciplinary Journal* 2, p. 127-160.
- JORM, A. F., SHARE, D. L., MCLEAN, R. & MATTHEWS, R. G. (1984). « Phonological recoding skill and learning to read: A longitudinal study ». *Applied Psycholinguistics* 5, p. 201-207.
- JUEL, C. (1988). « Learning to read and write: A longitudinal study of 54 children from first through fourth grade ». *Journal of Educational Psychology* 80, p. 437-447.
- KEENAN, J. M., BETJEMANN, R. S. & OLSON, R. K. (2008). « Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension ». *Scientific Studies of Reading* 12(3), p. 281-300.

- LÉTÉ, B., SPRENGER-CHAROLLES, L. & COLÉ, P. (2004). « MANULEX: A lexical database from French readers ». *Behavioral Research Methods, Instruments and Computers*, 36(1), p. 156-166.
- MELBY-LERVÅG, M., LYSTER, S. A. & HULME, C. (2012). « Phonological skills and their role in learning to read: A meta-analytic review ». *Psychological Bulletin* 138(2), p. 322-352.
- OAKHILL, J., CAIN, K. & BRYANT, P. (2003). « The dissociation between word reading and text comprehension: Evidence from component skills ». *Language and Cognitive Processes* 18(4), p. 443-466.
- O'REGAN, K. & LEVY-SCHOEN, A. (1978). « Les mouvements oculaires au cours de la lecture ». *L'Année psychologique* 78(2), p. 459-471.
- PACTON, S. (2008). « L'apprentissage de l'orthographe du français ». In : Desrochers, A., Martineau, F. & Morin, Y. C (éds), *Normes et pratiques orthographiques*. Ottawa : Éditions David, p. 331-354.
- PARVIAINEN, T., HELENIUS, P., POSKIPARTA, E. et al. (2006). « Cortical sequence of word perception in beginning readers ». *Journal of Neuroscience* 26(22), p. 6052-6061.
- PEEREMAN, R., SPRENGER-CHAROLLES, L. & MESSAOUD-GALUSI, S. (2013). « The contribution of morphology to the consistency of spelling-to-sound relations: A quantitative analysis based on French elementary school readers ». *L'Année psychologique* 113(1), p. 3-33.
- PERFETTI, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- PERFETTI, C. A., GOLDMAN, S. R. & HOGABOAM, T. W. (1979). « Reading skill and the identification of words in discourse context ». *Memory and Cognition* 7, p. 273-282.
- PIQUARD-KIPFFER, A. & SPRENGER-CHAROLLES, L. (2013). « Early predictors of future reading skills: A follow-up of French-speaking children from the beginning of kindergarten to the end of the second grade (age 5 to 8) ». *L'Année psychologique* 113(4), p. 491-521.
- SEYMOUR, P. H., ARO, M., & ERSKINE, J. M. (2003). « Foundation literacy acquisition in European orthographies ». *British Journal of Psychology* 94, p. 143-174.
- SHARE, D. L. (1995). « Phonological recoding and self-teaching: *Sine qua non* of reading acquisition ». *Cognition* 55, p. 151-218.
- SPENCER, M., QUINN, J. M. & WAGNER, R. K. (2014). « Specific reading disability: Major problem, myth or misnomer? ». *Learning Disabilities Research and Practice* 29(1), p. 3-9.
- SPRENGER-CHAROLLES, L. (1986). « Rôle du contexte linguistique, des informations visuelles et phonologiques dans la lecture et son apprentissage ». *Pratiques* 52, p. 9-27.
- SPRENGER-CHAROLLES, L. & COLÉ, P. (2013). *Lecture et Dyslexie. Approches cognitives*. Paris : Dunod.
- SPRENGER-CHAROLLES, L., SIEGEL, L. S. & BÉCHENNEC, D. (1998). « Phonological mediation and semantic and orthographic factors in silent reading in French ». *Scientific Studies of Reading* 2, p. 3-29.
- SPRENGER-CHAROLLES, L., SIEGEL, L. S., BÉCHENNEC, D. & SERNICLAES, W. (2003). « Development of phonological and orthographic processing in reading aloud, in silent reading and in spelling: A four year longitudinal study ». *Journal of Experimental Child Psychology* 84, p. 194-217.
- SPRENGER-CHAROLLES, L., SIEGEL, L. S. & BONNET, P. (1998). « Phonological mediation and orthographic factors in reading and spelling ». *Journal of Experimental Child Psychology* 68, p. 134-155.
- STANOVICH, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.

- STARR, M. S. & RAYNER, K. (2001). « Eye movements during reading: Some current controversies ». *Trends in Cognitive Sciences* 5, p. 156-163.
- STOTHARD, S. & HULME, C. (1992). « Reading comprehension difficulties in children ». *Reading and Writing: An Interdisciplinary Journal* 4(3), p. 245-256.
- WIMMER, H. & GOSWAMI, U. (1994). « The influence of orthographic consistency on reading development: Word recognition in English and German children ». *Cognition* 51, p. 91-103.
- ZIEGLER, J. C. & GOSWAMI, U. (2005). « Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory ». *Psychological Bulletin* 131(1), p. 3-29.
- ZIEGLER, J. C., PERRY, C. & ZORZI, M. (2014). « Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia ». *Philosophical Transactions of the Royal Society B* 369. En ligne : <http://dx.doi.org/10.1098/rstb.2012.0397>.

NOTES

1. L'écart-type (ET) est une mesure de la dispersion d'un ensemble de valeurs autour de leur moyenne. Plus il est petit, plus la population est homogène. Dans une courbe gaussienne normale, environ 70 % de la population est supposée avoir des scores entre -1 ET et +1 ET de la moyenne du groupe, et 15 % des scores soit à +1 ET, soit à -1 ET de cette moyenne.
2. Le terme *procédure* est plus adéquat (celui de *stratégie* laisse penser qu'il y a un choix).
3. Enfants scolarisés aux USA (le grade 1 correspond au début de l'apprentissage de la lecture, 6 ans).
4. D'autres études, qui n'ont pas été examinées dans cet article, indiquent que, quelle que soit la langue, les enfants utilisent des unités morphologiques quand ils lisent. Toutefois, cette capacité est sous la dépendance de facteurs phonologiques jusqu'à une période tardive (cf. Sprenger-Charolles & Colé, 2013, p. 102-109).
5. À la différence d'une étude longitudinale, dans une étude transversale ce sont des enfants différents qui sont examinés à chaque session.
6. Ce n'est pas le cas pour l'hypothèse d'un apprentissage par régularités statistiques non gouvernées par la phonologie qui a été développée par S. Pacton, P. Perruchet et M. Fayol (voir pour des synthèses en français, Pacton, 2008 et Sprenger-Charolles & Colé, 2013, p. 94-97).
7. Pour des études avec des enfants de langue française, voir J.-M. Braibant & F.-M. Gérard, 1996 et R. Goigoux, 2000.

RÉSUMÉS

Le présent article dresse un bilan de 30 ans de recherches sur l'apprentissage de la lecture. La première partie présente des résultats sur trois points essentiels à la compréhension des mécanismes en jeu dans la lecture et son apprentissage : 1) les capacités expliquant la compréhension écrite ; 2) le rôle joué par le contexte dans l'identification des mots écrits ; 3) les indicateurs de la mise en place d'automatismes dans cette identification. La seconde partie porte

sur l'incidence de la transparence de l'orthographe et le rôle de la phonologie dans le développement des capacités d'identification des mots écrits. La conclusion comporte un résumé des acquis de la recherche accompagné par un cadre explicatif, ainsi que les principales implications pour les politiques éducatives.

The present paper provides a review of 30 years of research on learning to read. The first section presents some results on three key issues to understand the mechanisms involved in learning to read: 1) Abilities explaining reading comprehension; 2) Context effects in word identification; 3) Automaticity in word identification. The second part deals with the influence of orthographic transparency in the development of word identification skills, and the role of phonological skills in this development. The conclusion provides an explanatory framework for research findings and some implications for education policies.

INDEX

Keywords : learning to read, reading comprehension, listening comprehension, decoding, written-word identification, automaticity in written-word identification, grapheme-phoneme correspondences, French orthography

Mots-clés : apprentissage de la lecture, compréhension écrite, compréhension orale, décodage, identification des mots écrits, automatismes dans l'identification des mots écrits, correspondances graphème-phonème, orthographe du français

AUTEUR

LILIANE SPRENGER-CHAROLLES

Aix-Marseille Université, Centre national de la recherche scientifique, LPC, UMR 7290, F-13331, France